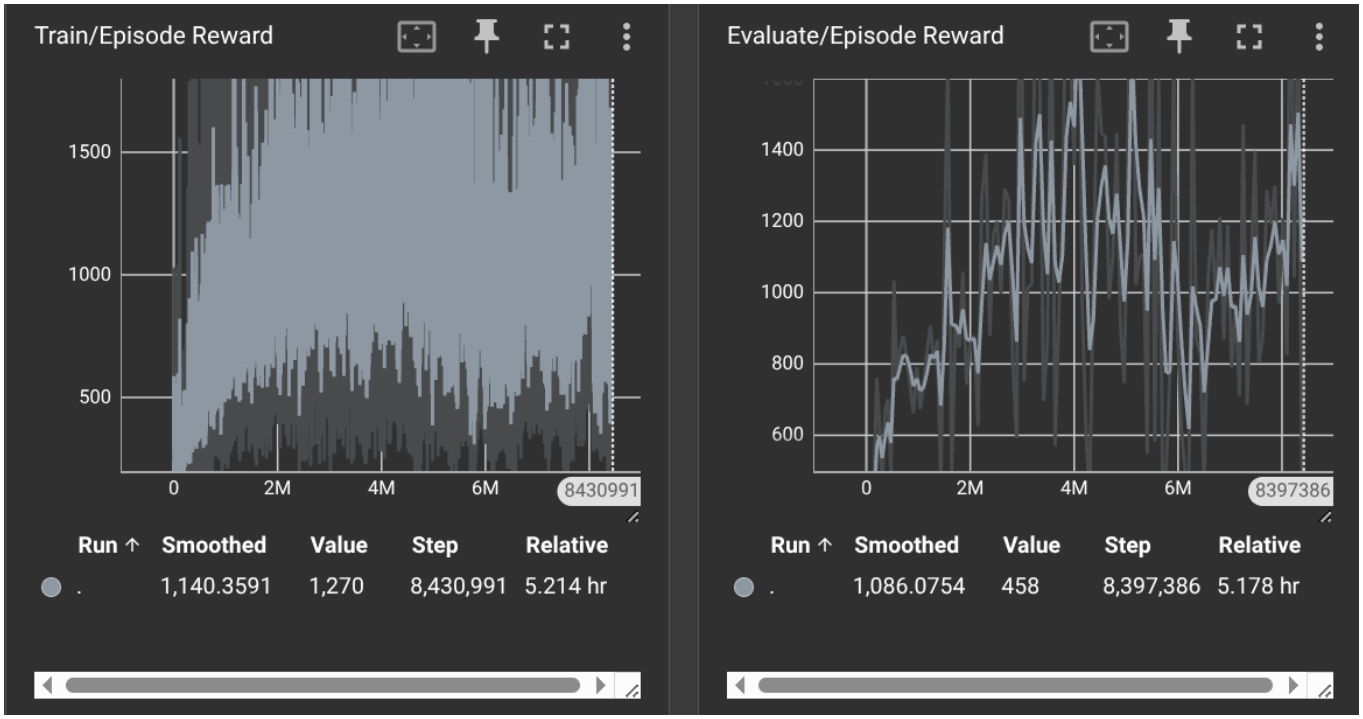
# Atari DQN HW2

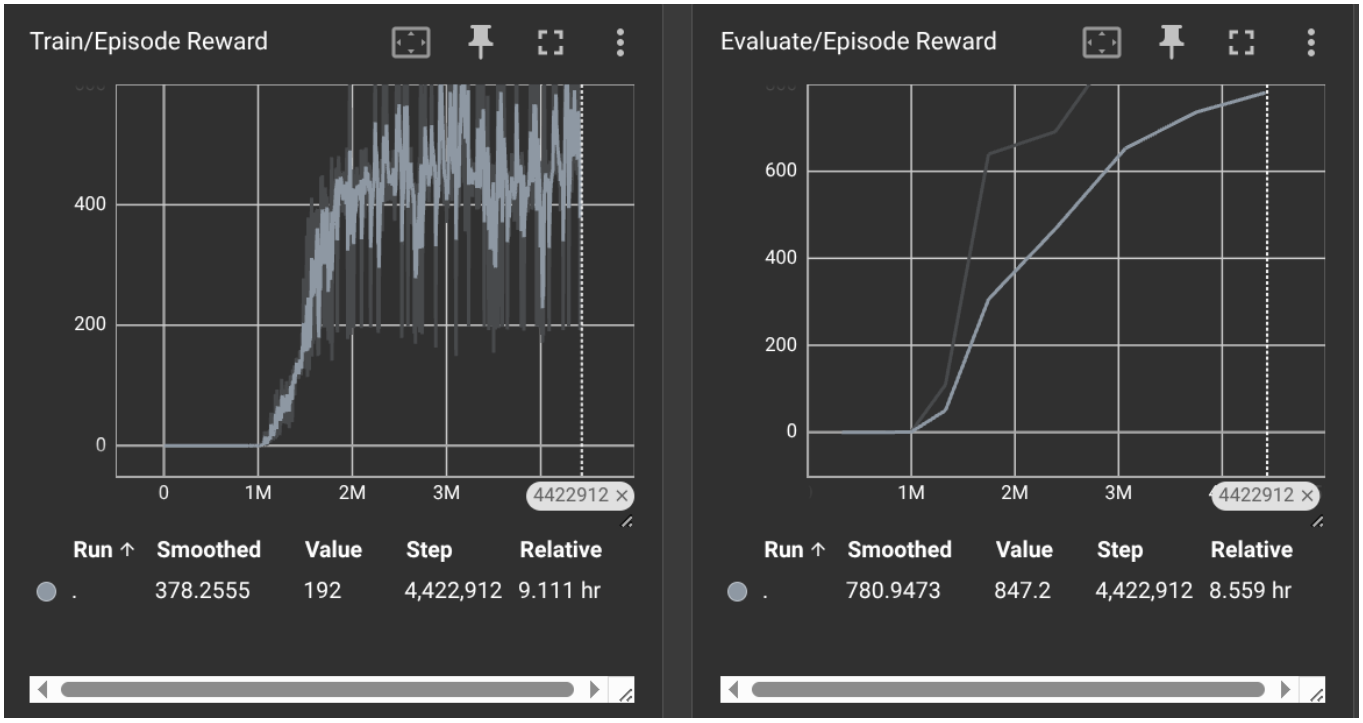
## Report

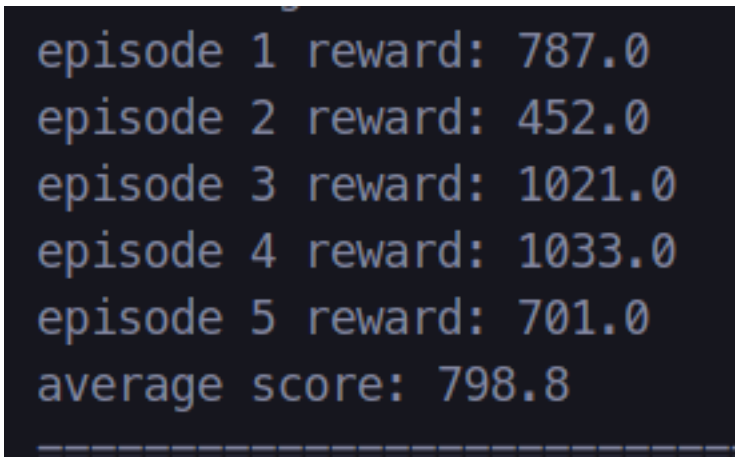
- Screenshot of Tensorboard training curve and testing results on DQN



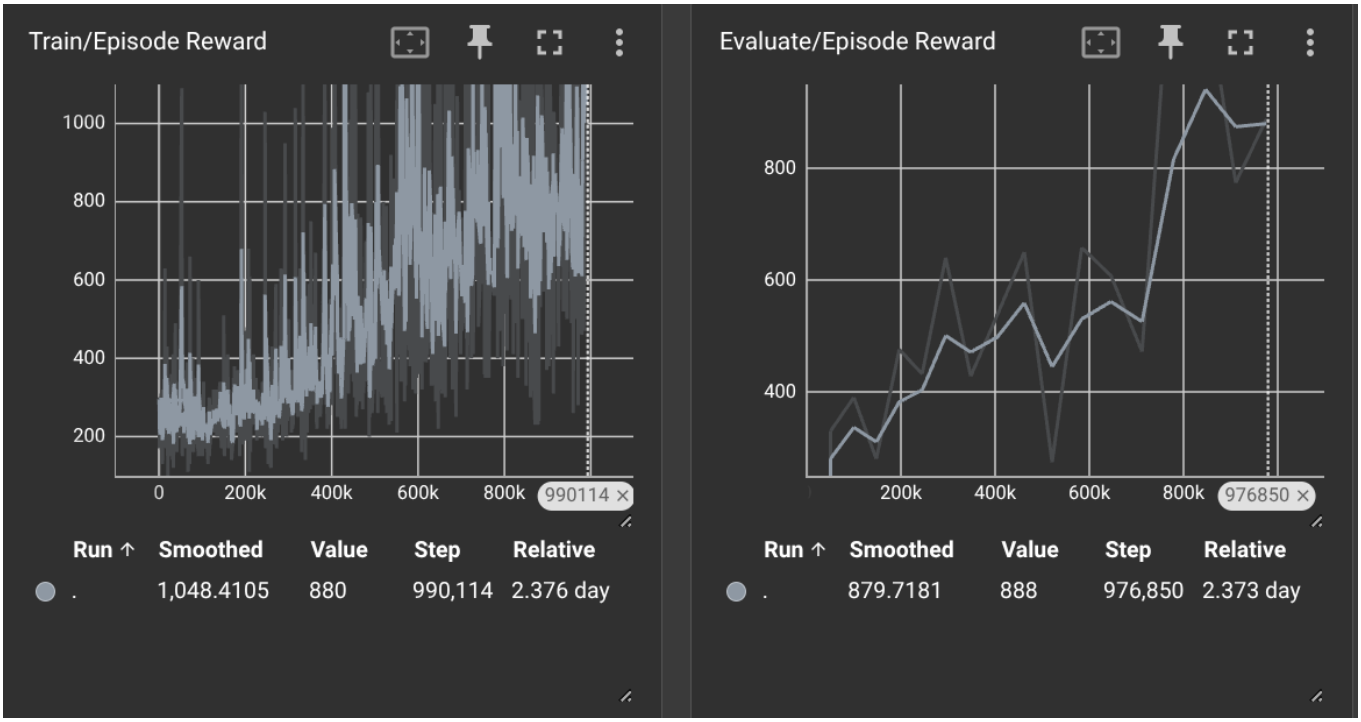

```
episode 1 reward: 1400.0
episode 2 reward: 1150.0
episode 3 reward: 730.0
episode 4 reward: 880.0
episode 5 reward: 1200.0
average score: 1072.0
```

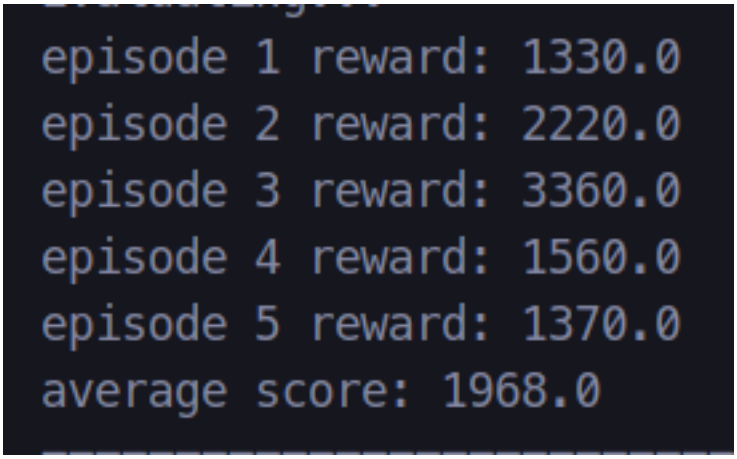
- Screenshot of Tensorboard training curve and testing results on Enduro-v5 using DQN




- Screenshot of Tensorboard training curve and testing results on DDQN, and discuss the difference between DQN and DDQN

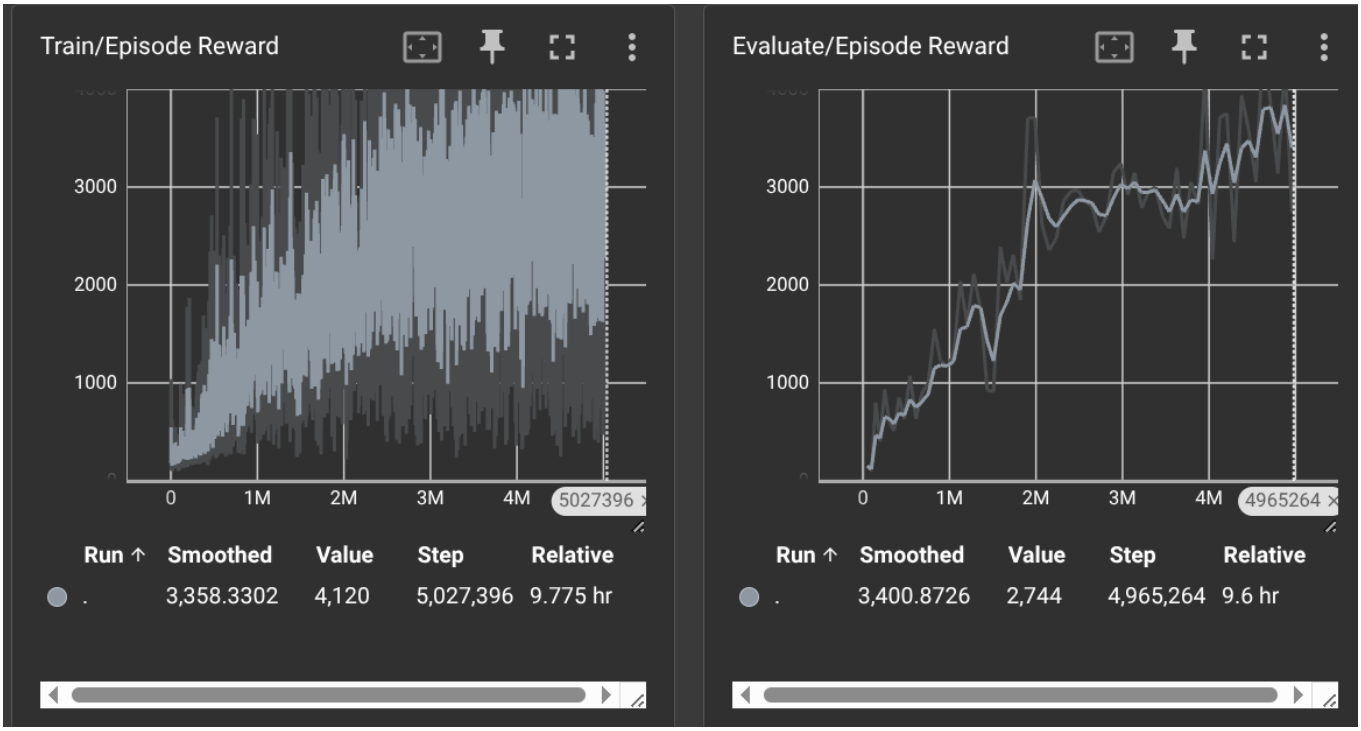


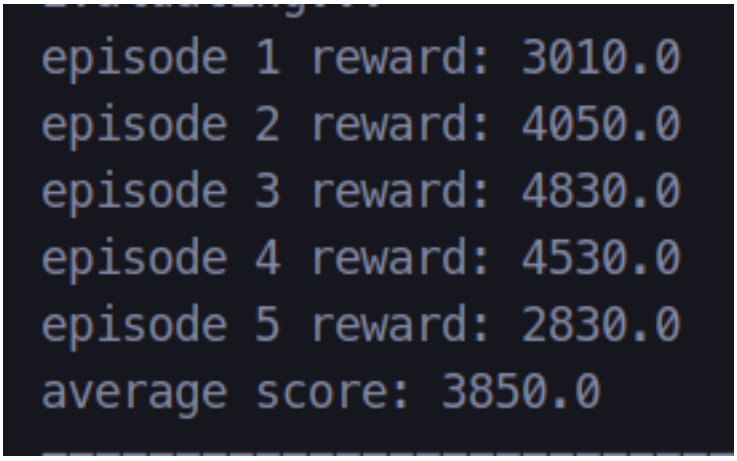




– Discussion

DQN overestimates Q-values since one network selects and evaluates actions. DDQN separates these roles, reducing bias and yielding smoother, more stable learning.

- Screenshot of Tensorboard training curve and testing results on Dueling DQN, and discuss the difference between DQN and Dueling DQN

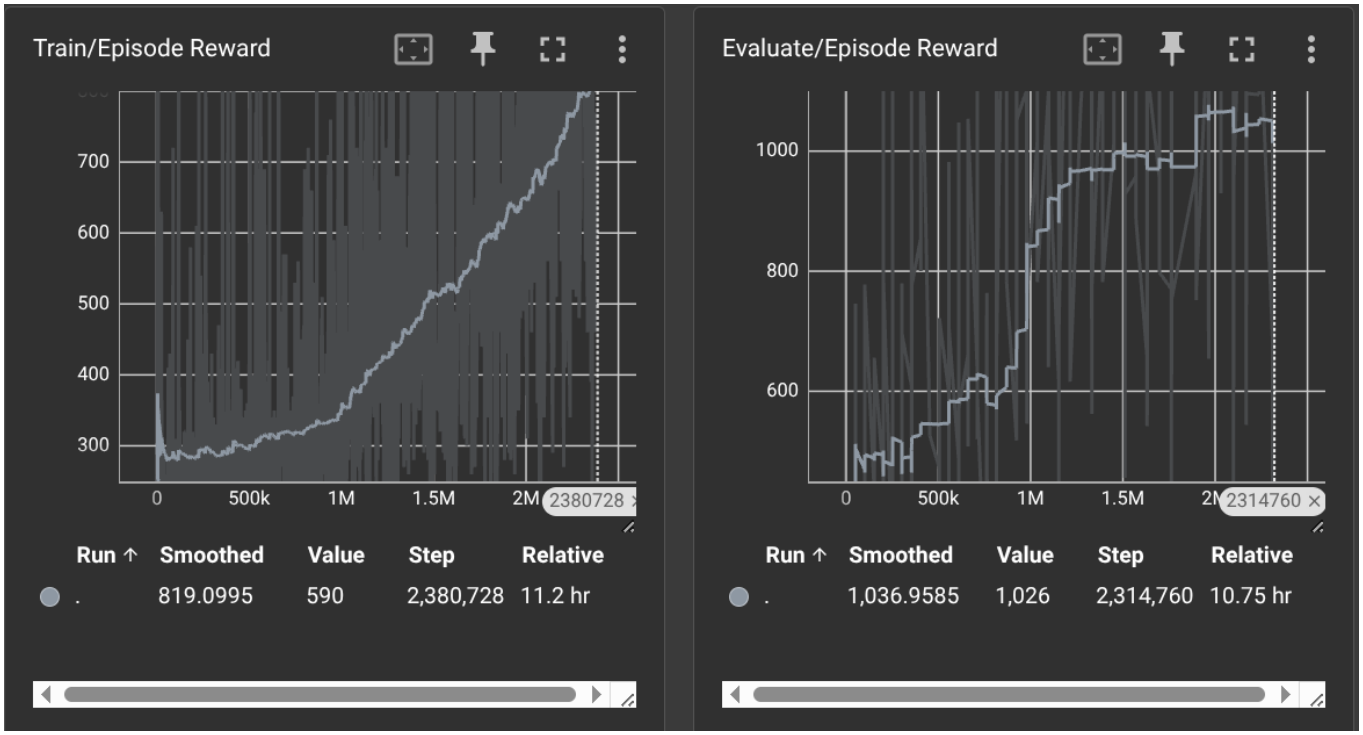


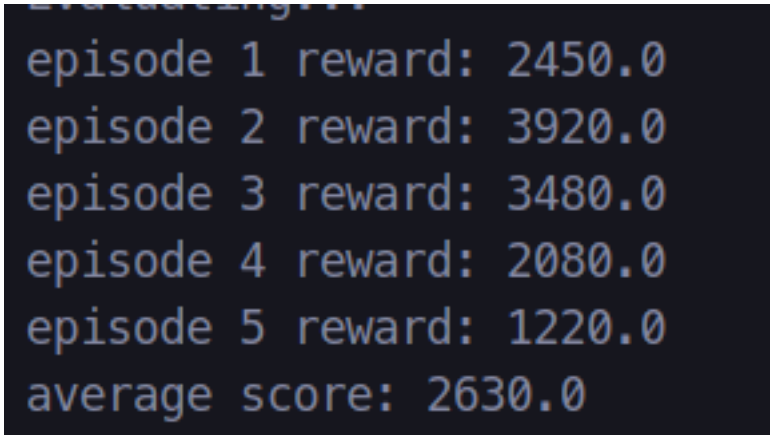

– Discussion

Dueling DQN splits Q-values into state-value and advantage, improving representation and efficiency. It outperforms vanilla DQN, especially where many actions are irrelevant.

- Screenshot of Tensorboard training curve and testing results on DQN with parallelized rollout, and discuss the difference between DQN and DQN with parallelized rollout




for max smooth




– Discussion

Parallelized rollout runs multiple environments, giving diverse, uncorrelated samples. This speeds training and yields smoother, more stable reward curves than standard DQN.